

CBM

R

7626

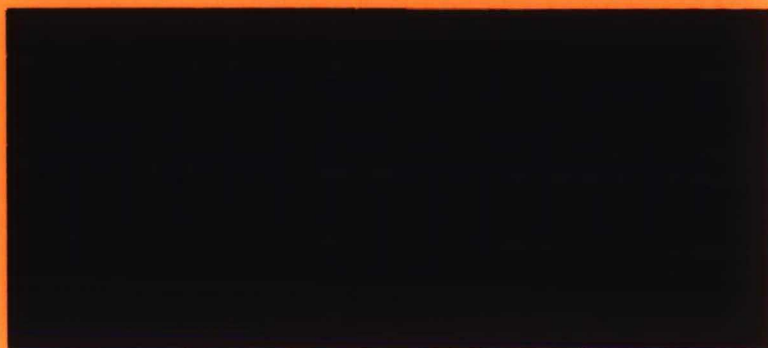
1986

212



faculteit der economische wetenschappen

## RESEARCH MEMORANDUM



TILBURG UNIVERSITY

DEPARTMENT OF ECONOMICS

Postbus 90153 - 5000 LE Tilburg

Netherlands





EXPERIMENTAL DESIGN AND REGRESSION ANALYSIS  
IN SIMULATION:

AN FMS CASE STUDY

Jack P.C. Kleijnen<sup>1)</sup>  
and  
Charles R. Standridge<sup>2)</sup>  
(Senior Member)

lg  
510.5  
510.9

- 1) Department of Information Systems and Accountancy,  
School of Business and Economics, Catholic University  
Tilburg (Katholieke Hogeschool Tilburg, P.O. Box 90153,  
5000 LE Tilburg, Netherlands.
- 2) Pritsker & Associates, Inc., 5412 88th Street, Lubbock,  
Texas 79424.

## ABSTRACT

The machine mix for a particular FMS, the number of machines performing each of three operations and the number of machines performing any of the three operations (flexible machines), is input to an FMS simulation. An intuitively selected combination of these four inputs are compared to a  $2^{4-1}$  fractional factorial design. The throughput predicted by the simulation is analyzed through two different regression models. These models are validated. A regression model in two inputs including their interaction, gives valid predictions and stable explanations.

Key Words: statistics, fractional factorial design, regression model, Analysis of Variance model, metamodel, validation, cross-validation, flexible manufacturing system, case study.

## INTRODUCTION

Simulation is a technique applied in many areas because of its flexibility, simplicity and realism. However, because simulation involves experimenting (with the model of a real system) it requires statistical design and analysis. The present paper concentrates on strategic issues, namely, which variants of the simulation model are actually run (i.e., which combinations of parameter values are input), and how can the resulting output be analyzed? Strategic issues arise in both random and deterministic simulation, whereas tactical issues (like runlength and confidence intervals) arise only in random simulation. The case-study of the present paper concerns a deterministic simulation model of a Flexible Manufacturing System (FMS).



## FMS

Figure 1 shows one possible layout of the flexible manufacturing system being studied [2]. A cart path surrounds the machines. One cart moves parts of one type between the wash station and the machine to perform the next operation. After the operation is completed, the cart moves the part back to the wash station. Parts are washed before each operation is performed and after the last operation. Parts enter from the lathes and exit to the inspection station. There are three operations, OP.10, OP.20 and OP.30. In Figure 1, five machines perform OP.10, two OP.20, two OP.30 and one is flexible to perform any of the three operations.

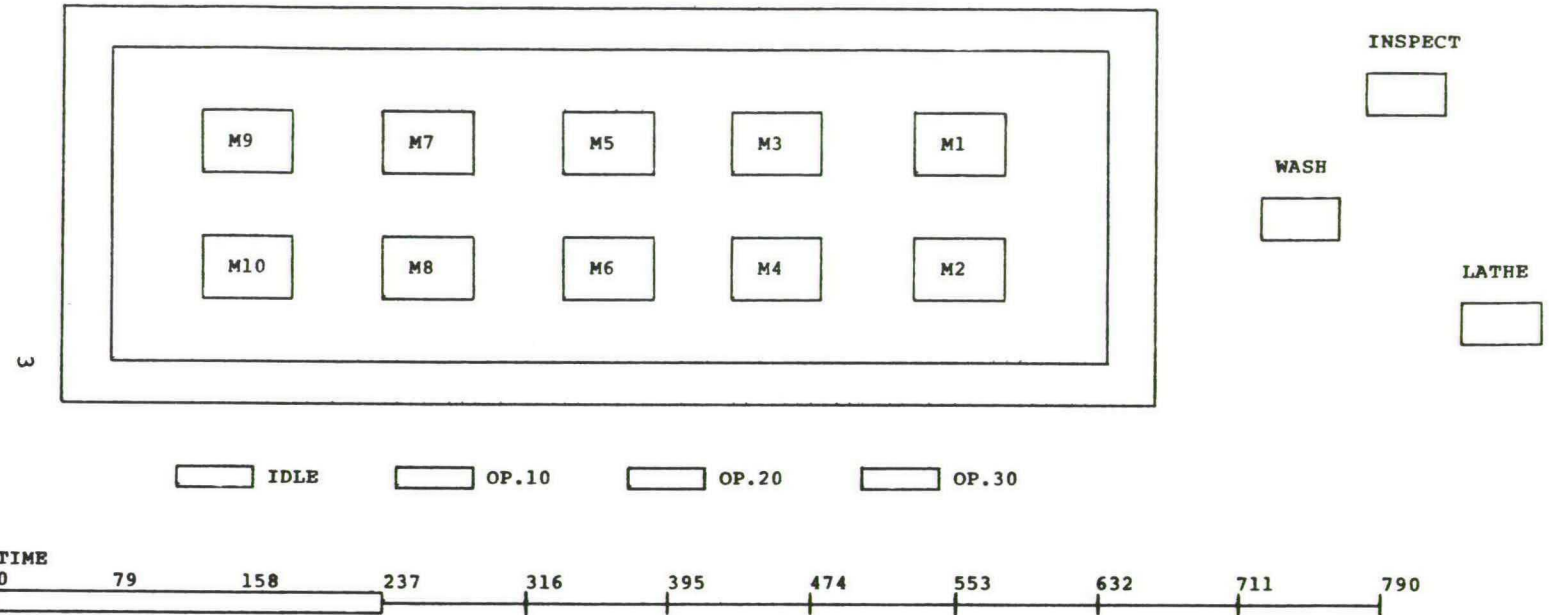
## PROBLEM SCOPE

The case-study determines the machine mix, the number of machines performing each operation and the number of flexible machines performing any operation:

$x_1$ :	number of machines for operation 10	
$x_2$ :	"	20
$x_3$ :	"	30
$x_4$ :	number of flexible machines for operation 10, 20 or 30.	

The machine mix (combination of  $x_1, x_2, x_3, x_4$ ) should produce a throughput of 3140 parts per week. The scope of the case-study is restricted to a given control logic of the FMS. The system operates on only one part type.

# EVALUATION OF FMS DESIGN



An algebraic, naive solution to the machine-mix problem is typically computed as follows. The throughput constraint (3140 parts per week) means that the number of parts produced per minute is 0.73. It is known that operations 10, 20 and 30 require per part: 7.5, 2.5 and 3.5 minutes respectively. Hence the computed machine mix is  $x_1 = 5.4$ ,  $x_2 = 1.8$  and  $x_3 = 2.5$ .

Rounding these results up produces a machine mix of  $x_1 = 6$ ,  $x_2 = 2$ , and  $x_3 = 3$ . This machine mix provides excess capacity for each machine type. Thus no flexible machines are necessary,  $x_4 = 0$ . Rounding these results down produces a machine mix of  $x_1 = 5$ ,  $x_2 = 1$  and  $x_3 = 2$ . The capacity needed from flexible machines is  $(0.4 + 0.8 + 0.5 = 1.7)$ . Thus  $x_4 = 2$ . Flexible machines are much more expensive than machines performing one operation (fixed machines). Thus, substituting flexible machines for fixed machines need not be considered.

The naive solution suggests the following experimental area over which the inputs to the simulation may vary:

$$\begin{aligned} 5 &\leq x_1 \leq 6 \\ 1 &\leq x_2 \leq 2 \\ 2 &\leq x_3 \leq 3 \\ 0 &\leq x_4 \leq 2 \end{aligned}$$

#### EXPERIMENTAL DESIGN

Given the problem of the preceding section, 24 combinations might be simulated. In many simulation studies the number of combinations is much greater! How a fraction of all possible combinations still yields adequate results is illustrated.

Table 1 specifies an intuitively selected fraction of input combinations, specified without a knowledge of the statistical theory of experimental design.

Table 1. Intuitive fractional design <sup>1)</sup>

Run i	Operat. 10 $x_1$	Operat. 20 $x_2$	Operat. 30 $x_3$	Flexible $x_4$	(Total $\Sigma x_j$ )
1	-	-	-	-	(8)
2	+	-	+	M	(11)
3	-	+	+	M	(11)
4	+	+	-	-	(10)
5	-	-	+	+	(11)
6	+	-	-	M	(10)
7	-	+	-	M	(10)
8	+	+	+	+	(13)

1) Legend: - minimum value of  $x_j$  ( $j = 1, \dots, 4$ )  
 + maximum "  
 M middle value of  $x_4$  ( $x_4 = 1$ ).

Formal experimental design theory proceeds as follows. Each "factor" (independent variable  $x_j$  with  $j = 1, \dots, 4$ ) is studied at only two "levels" (values). (In the case-study the factors have a very narrow range and the levels must be integer numbers; hence the two levels coincide with the possible levels for the first three factors; obviously a possible value for  $x_4$  is one, which does not coincide with its levels, zero and two; see eq. 1.) This restriction still yields  $2^4$  combinations of input values. Statistical theory suggest that only a fraction of these  $2^4$  combinations be simulated. How big that fraction should be depends on the (regression or Analysis of Variance) metamodel. If it could be assumed that the simulation model yields a throughput  $y$  equal to the additive effects of the four inputs  $x_j$  ( $j = 1, \dots, 4$ ), then only four input combinations would suffice. In other words, if the relationship between the simulation model's output  $y$  and its inputs  $x_j$  could be approximated by the metamodel



$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3 + \hat{\beta}_4 x_4 \quad (2)$$

then the four estimated effects  $\hat{\beta}_j$  could be computed from only four simulation runs. Actually it seems dangerous to assume apriori that the simulation inputs do not interact. Therefore more than four combinations are simulated. Statistical theory (see Kleijnen, 1986) results in the fractional design of Table 2.

Table 2. Formal  $2^{4-1}$  fractional factorial design

Run	Operat. 1	Operat. 2	Operat. 3	Flexible	(Total
i	$x_1$	$x_2$	$x_3^{1)}$	$x_4$	$\Sigma x_j$ )
1	-	-	-	-	(8)
2	+	-	+	-	(10)
3	-	+	+	-	(10)
4	+	+	-	-	(10)
5	-	-	+	+	(11)
6	+	-	-	+	(11)
7	-	+	-	+	(11)
8	+	+	+	+	(13)

1)  $x_{i3} = x_{i1} x_{i2} x_{i4}$  ( $i = 1, \dots, 8$ ); also see [1].

A "run" means that a row of Table 1 or 2 is translated into the input values  $x_{ij}$ ; this input is transformed by the simulation model into the simulation response  $y_i$ . Both the intuitive and the formal design require eight simulation runs:  $n = 8$ . The only difference between the formal and intuitive design is the use of  $x_4$  equals to 1 in runs 2, 3, 6 and 7. The next section will show that the formal design gives better conclusions.

100

100

100

100

Although the experimental designs of Tables 1 and 2 comprise more than four runs, the analysis of the experimental results starts assuming that the simple additive regression model of eq. (2) is valid; this assumption will be checked later. The effects  $\hat{\beta}_j$  of eq. (2) are computed (the metamodel is "calibrated") using the Ordinary Least Squares (OLS) algorithm

$$\hat{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{y} \quad (3)$$

where  $\hat{\beta}' = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4)$ ,  $\tilde{X} = (x_{ij})$  with  $i = 1, \dots, 8$  and  $j = 1, \dots, 4$ ,  $\tilde{y}' = (y_1, \dots, y_8)$ . the standard OLS algorithm also yields the covariance matrix

$$\hat{\Omega}_{\hat{\beta}} = (\tilde{X}'\tilde{X})^{-1}\sigma^2 \quad (4)$$

where  $\sigma^2$  denotes the variance of the error terms  $y - \hat{y}$ , that is, if all possible combinations of inputs  $x_j$  were simulated resulting in the simulation responses  $y$  and if the inputs  $x_j$  were also used in the regression model of eq. (2), then the prediction errors  $e = y - \hat{y}$  would form a statistical distribution with variance  $\sigma^2$ ; see Kleijnen (1986). This  $\sigma^2$  is estimated through the Mean Squared Residuals (MSR):

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-q} \quad (5)$$

where  $q$  denotes the number of effects ( $q=5$ ). Obviously  $\hat{\text{var}}(\hat{\beta}_j)$  is the  $j^{\text{th}}$  element on the main diagonal of  $\hat{\Omega}_{\hat{\beta}}$  obtained from eqs. (4) and (5). Eqs. (2) through (5) yield Table 3, which shows that the formal design gives more accurate estimates of the factor effects. Therefore the rest of the paper concentrates on the results of the formal design.

Table 3. Estimated variance of estimated effects in additive model:  $\hat{\text{var}}(\hat{\beta}_j)$ .

<u>Effect</u>	<u>Intuitive design</u>	<u>Formal design</u>
$\hat{\beta}_1$ (operation 1)	0.5	0.5
$\hat{\beta}_2$ (operation 2)	0.5	0.5
$\hat{\beta}_3$ (operation 3)	1.0	0.5
$\hat{\beta}_4$ (flexible)	0.5	0.13
$\hat{\beta}_0$ (constant)	20.6	19.6

One way of validating the calibrated metamodel would be to ask: what happens to the values of the estimated effects  $\hat{\beta}_j$  if one run is deleted from Table 2? Obviously these values do change, certainly the value of the non-significant effects. This significance is tested through the t-statistic:

$$t_v^{(j)} = \frac{\hat{\beta}_j}{\hat{\text{var}}(\hat{\beta}_j)} \quad j = 1, \dots, 4 \quad (6)$$

where  $v = n - q$  so that  $v = 3$  if no run is deleted and  $v = 2$  if one run is deleted. The significant effects should remain stable upon run deletion. Table 4 displays only the effects that are significant at  $\alpha = 0.3$ .

A more compact way of evaluating the effect of run deletion, is to concentrate on the predictor  $\hat{y}$ , in other words, the criterion becomes prediction instead of explanation. Table 5 displays the "relative prediction error"

$$r_i = \frac{y_i - \hat{y}_i}{\hat{y}_i} \quad (i = 1, \dots, 8) \quad (7)$$



2012 3 9

2012 3 9

2

2012 3 9

2

2012 3 9

2012 3 9

2012 3 9

2012 3 9

2012 3 9

2012 3 9

2012 3 9

2012 3 9

where  $\hat{y}_i$  denotes the predicted response (throughput) using eq. (2) but now estimating the effects  $\hat{\beta}_j$  using only  $n-1$  runs, namely deleting run  $i$  from the simulation data: cross-validation.

Table 4. Stability of significant  $\hat{\beta}$  upon run deletion.

Run deleted	$\hat{\beta}_1$ (op.10)	$\hat{\beta}_2$ (op.20)	$\hat{\beta}_3$ (op.30)	$\hat{\beta}_4$ (flex.)	$\hat{\beta}_0$ (const.)
1		557		577	
2		712		500	2032
3		640		700	
4		629		694	
5				658	1962
6				736	
7				536	
8				541	3288
none				541	3288

Table 5. Cross-validation:  $(y_i - \hat{y}_i)/y_i$ .

Run deleted:	1	2	3	4	5	6	7	8
Relative error: (in %)	10	27	-19	-18	13	33	-38	-35

The errors in Table 5 and the instabilities in Table 4 are so large that the additive regression model of eq. (2) is rejected, and a new metamodel is investigated (using the old data of the simulation experiment).

# AN ALTERNATIVE METAMODEL

Table 4 suggests that the factors  $x_1$  and  $x_3$  are not important. Therefore a regression model in the remaining factors ( $x_2$  and  $x_4$ ) is formulated, including possible interaction between these two factors:

$$y = \hat{\gamma}_0 = \hat{\gamma}_2 x_2 + \hat{\gamma}_4 x_4 + \hat{\gamma}_{2,4} x_2 x_4 \quad (8)$$

Table 4 is now replaced by Table 6: in the alternative model all estimated effects remain significant upon deletion of run  $i$  ( $i = 1, \dots, 8$ ).

Table 6. Stability of significant  $\hat{\gamma}$  upon run deletion

Run deleted	$\hat{\gamma}_2$ (op.20)	$\hat{\gamma}_4$ (flex.)	$\hat{\gamma}_{2,4}$ (interact.)	$\hat{\gamma}_0$ (const.)
1	952	1364	-492	776
2	952	1300	-460	776
3	952	1324	-468	776
4	952	1340	-484	776
5	1152	1432	-576	576
6	752	1232	-376	976
7	952	1332	-476	776
8	952	1332	-476	776
none	952	1332	-476	776

Table 5 is now replaced by Table 7. The relative prediction errors become much smaller. The model can also be double-checked in this case-study, because there are simulation data available for some extra input combinations (besides the eight runs of Table 2), namely the combinations listed in Table 1 but not in Table 2. The four combinations in Table 8 were not used to calibrate eq. (8) and are now used to double-check the predictive power of the alternative regression (meta)model.

Table 7. Cross-validation of eq. (8)

Run deleted:	1	2	3	4	5	6	7	8
Rel. error(%)	2	2	-1	1	-16	14	0	0

Table 8. Double-check of eq. (8).

$x_2$ (op.20)	$x_4$ (flex.)	$x_2x_4$ (interact.)	$\hat{y}$ (thruput)	(predict.)	r(%) (error)
2	1	2	1368	3060	3
1	1	1	3456	2584	25
2	1	2	3408	3060	10
1	1	1	2896	2584	11

Thus the meta-model is

$$\hat{y} = 776 + 952x_2 + 1334x_4 - 476x_2x_4$$



This model suggests:

1. Machines performing operation 20 and flexible machines are the bottlenecks in the system.
2. There is a trade-off between using more machines performing operation 20 and more flexible machines as shown by the negative coefficient in the interaction term.

### CONCLUSION

The present case-study illustrates how experimental design and regression analysis can be applied to evaluate an FMS. The statistical techniques are quite simple, i.e., the FMS simulation model was available at the beginning of the study and within a few days the results of the present paper (design and analysis) were obtained. Statistical techniques are not an aim in themselves. They can reduce the drawbacks of an empirical technique like simulation, i.e., at the end of this quick-and-dirty case-study the regression metamodel of eq. (8) helped the authors to better understand how an FMS works!

### REFERENCES

1. Kleijnen, J.P.C. (1986), Statistical Tools for Simulation Practitioners, Marcel Dekker, New York.
2. Wortman, David B. and James R. Wilson, "Optimizing a Manufacturing Plant by Computer Simulation", Computer-Aided Engineering, September, 1984.

IN 1985 REEDS VERSCHENEN

- 168 T.M. Doup, A.J.J. Talman  
A continuous deformation algorithm on the product space of unit  
simplices
- 169 P.A. Bekker  
A note on the identification of restricted factor loading matrices
- 170 J.H.M. Donders, A.M. van Nunen  
Economische politiek in een twee-sectoren-model
- 171 L.H.M. Bosch, W.A.M. de Lange  
Shift work in health care
- 172 B.B. van der Genugten  
Asymptotic Normality of Least Squares Estimators in Autoregressive  
Linear Regression Models
- 173 R.J. de Groof  
Geïsoleerde versus gecoördineerde economische politiek in een twee-  
regiomodel
- 174 G. van der Laan, A.J.J. Talman  
Adjustment processes for finding economic equilibria
- 175 B.R. Meijboom  
Horizontal mixed decomposition
- 176 F. van der Ploeg, A.J. de Zeeuw  
Non-cooperative strategies for dynamic policy games and the problem  
of time inconsistency: a comment
- 177 B.R. Meijboom  
A two-level planning procedure with respect to make-or-buy deci-  
sions, including cost allocations
- 178 N.J. de Beer  
Voorspelprestaties van het Centraal Planbureau in de periode 1953  
t/m 1980
- 178a N.J. de Beer  
BIJLAGEN bij Voorspelprestaties van het Centraal Planbureau in de  
periode 1953 t/m 1980
- 179 R.J.M. Alessie, A. Kapteyn, W.H.J. de Freytas  
De invloed van demografische factoren en inkomen op consumptieve  
uitgaven
- 180 P. Kooreman, A. Kapteyn  
Estimation of a game theoretic model of household labor supply
- 181 A.J. de Zeeuw, A.C. Meijdam  
On Expectations, Information and Dynamic Game Equilibria

- 182 Cristina Pennavaja  
Periodization approaches of capitalist development.  
A critical survey
- 183 J.P.C. Kleijnen, G.L.J. Kloppenburg and F.L. Meeuwssen  
Testing the mean of an asymmetric population: Johnson's modified T  
test revisited
- 184 M.O. Nijkamp, A.M. van Nunen  
Freia versus Vintaf, een analyse
- 185 A.H.M. Gerards  
Homomorphisms of graphs to odd cycles
- 186 P. Bekker, A. Kapteyn, T. Wansbeek  
Consistent sets of estimates for regressions with correlated or  
uncorrelated measurement errors in arbitrary subsets of all  
variables
- 187 P. Bekker, J. de Leeuw  
The rank of reduced dispersion matrices
- 188 A.J. de Zeeuw, F. van der Ploeg  
Consistency of conjectures and reactions: a critique
- 189 E.N. Kertzman  
Belastingstructuur en privatisering
- 190 J.P.C. Kleijnen  
Simulation with too many factors: review of random and group-  
screening designs
- 191 J.P.C. Kleijnen  
A Scenario for Sequential Experimentation
- 192 A. Dortmans  
De loonvergelijking  
Afwenteling van collectieve lasten door loontrekkers?
- 193 R. Heuts, J. van Lieshout, K. Baken  
The quality of some approximation formulas in a continuous review  
inventory model
- 194 J.P.C. Kleijnen  
Analyzing simulation experiments with common random numbers
- 195 P.M. Kort  
Optimal dynamic investment policy under financial restrictions and  
adjustment costs
- 196 A.H. van den Elzen, G. van der Laan, A.J.J. Talman  
Adjustment processes for finding equilibria on the simplotope

- 197 J.P.C. Kleijnen  
Variance heterogeneity in experimental design
- 198 J.P.C. Kleijnen  
Selecting random number seeds in practice
- 199 J.P.C. Kleijnen  
Regression analysis of simulation experiments: functional software specification
- 200 G. van der Laan and A.J.J. Talman  
An algorithm for the linear complementarity problem with upper and lower bounds
- 201 P. Kooreman  
Alternative specification tests for Tobit and related models



IN 1986 REEDS VERSCHENEN

- 202 J.H.F. Schilderlinck  
Interregional Structure of the European Community. Part III
- 203 Antoon van den Elzen and Dolf Talman  
A new strategy-adjustment process for computing a Nash equilibrium  
in a noncooperative more-person game
- 204 Jan Vingerhoets  
Fabrication of copper and copper semis in developing countries.  
A review of evidence and opportunities.
- 205 R. Heuts, J. v. Lieshout, K. Baken  
An inventory model: what is the influence of the shape of the lead  
time demand distribution?
- 206 A. v. Soest, P. Kooreman  
A Microeconomic Analysis of Vacation Behavior
- 207 F. Boekema, A. Nagelkerke  
Labour Relations, Networks, Job-creation and Regional Development  
A view to the consequences of technological change
- 208 R. Alessie, A. Kapteyn  
Habit Formation and Interdependent Preferences in the Almost Ideal  
Demand System
- 209 T. Wansbeek, A. Kapteyn  
Estimation of the error components model with incomplete panels
- 210 A.L. Hempenius  
The relation between dividends and profits
- 211 J. Kriens, J.Th. van Lieshout  
A generalisation and some properties of Markowitz' portfolio  
selection method

Bibliotheek K. U. Brabant



17 000 01059734 3